

ABSTRACT OF THE DISCLOSURE

A biological marker identification method identifies biological markers within broad sets of biological data containing many more measurements than observations. For example, the data can contain thousands of measurements on each blood sample obtained from fewer than 100 subjects, each of which falls into one of a set of clinical classes or is associated with a value of a continuous clinical response variable. At least one biomarker, containing a small subset of measurements, is found that is capable of predicting a clinical endpoint. The biomarker can be used for, e.g., diagnosing disease or assessing response to a drug. First, the set of measurements is reduced to a smaller set of candidate measurements by eliminating measurements that either cannot distinguish among classes or are redundant. Biomarker subsets are then selected from the remaining set of measurements, either by an exhaustive search or a heuristic method that finds good but not necessary globally optimal biomarkers.

S:\Client Folders\Surromed\79\UTILITY\SURR-79.doc